Problem 1. \textit{(Non-coherent Detection)}

1. We simply find the conditional density of \(r_0, \ldots, r_{N-1}\) under the two hypotheses and derive the likelihood ratio.

For \(H_0\), with pure noise assumption we have:

\[ P(r_0, \ldots, r_{N-1}|H_0) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left( -\sum_{i=0}^{N-1} \frac{r_i^2}{2\sigma^2} \right) \]

And for \(H_1\) we have:

\[
\begin{align*}
P(r_0, \ldots, r_{N_1}|H_1) &= \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=0}^{N-1} \left( r_i - A \cos(2\pi f_0 i) \right)^2 \right) \\
&= \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=0}^{N-1} r_i^2 + \frac{A}{\sigma^2} \sum_{i=0}^{N-1} \cos(2\pi f_0 i) r_i - \frac{NA^2}{2\sigma^2} - \frac{1}{N} \sum_{i=0}^{N-1} \cos^2(2\pi f_0 i) \right) \\
&\approx \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=0}^{N-1} r_i^2 + \frac{A}{\sigma^2} \sum_{i=0}^{N-1} \cos(2\pi f_0 i) r_i - \frac{NA^2}{4\sigma^2} \right)
\end{align*}
\]

Then we can compute the likelihood ratio:

\[
L(r_0, \ldots, r_{N-1}) = \frac{P(r_0, \ldots, r_{N-1}|H_1)}{P(r_0, \ldots, r_{N-1}|H_0)} = \exp \left( \frac{A}{\sigma^2} \sum_{i=0}^{N-1} r_i \cos(2\pi f_0 i) - \frac{NA^2}{4\sigma^2} \right) \leq \frac{P_0}{P_1} = 1,
\]

which gives us the following MAP rule:

\[
\sum_{i=0}^{N-1} r_i \cos(2\pi f_0 i) \leq \frac{NA}{H_1}.
\]

Now we can compute the conditional and mean error probabilities.
Under $H_0: \ r_i \sim N(0, \sigma^2)$ which implies that

$$\sum_{i=0}^{N-1} r_i \cos(2\pi f_0 i) \sim N \left( 0, \sigma^2 \sum_{i=0}^{N-1} \cos^2(2\pi f_0 i) \right) \approx N \left( 0, \frac{N\sigma^2}{2} \right).$$

Hence we obtain

$$P(E|H_0) = P \left( N \left( 0, \frac{N\sigma^2}{2} \right) > \frac{NA}{4} \right)$$

$$= Q \left( \frac{\frac{NA}{4}}{\sqrt{\frac{N\sigma^2}{2}}} \right)$$

$$= Q \left( \sqrt{\frac{N\sigma^2}{8\sigma^2}} \right) = Q \left( \sqrt{\frac{SNR}{8}} \right).$$

Similarly under $H_1: r_i \sim N(A \cos(2\pi f_0 i), \sigma^2)$. If we define $Z \overset{\text{def}}{=} \sum_{i=0}^{N-1} r_i \cos(2\pi f_0 i)$ then $Z$ is a Gaussian random variables and

$$E[Z] = \sum A \cos^2(2\pi f_0 i) \approx \frac{NA}{2}$$

$$Var[Z] = \sigma^2 \sum_{i=0}^{N-1} \cos^2(2\pi f_0 i) = \frac{N\sigma^2}{2}.$$ 

Hence $Z \sim N \left( \frac{NA}{2}, \frac{N\sigma^2}{2} \right)$. For error probability under $H_1$ we have

$$P(E|H_1) = P \left( Z < \frac{NA}{4} \right) = Q \left( \frac{\frac{NA}{2} - \frac{NA}{4}}{\sqrt{\frac{N\sigma^2}{2}}} \right) = Q \left( \sqrt{\frac{SNR}{8}} \right).$$

Combining the conditional error probabilities we obtain the mean error probability as

$$P(E) = \frac{1}{2} P(E|H_0) + \frac{1}{2} P(E|H_1) = Q \left( \sqrt{\frac{SNR}{8}} \right).$$

2. You may know from signal processing courses that $\sum_{i=0}^{N-1} r_i \cos(2\pi f_0 i)$ is simply the Discrete Cosine Transform of the sequence $(r_0, r_1, ..., r_{N-1})$. Actually the MAP rule tries to analyze the signal in the frequency domain. If it observes any peak of considerable height at frequency $f_0$, it chooses $H_1$, otherwise it chooses $H_0$, which intuitively seems correct.

3. The structure of the MAP rule does not change. In other words, the MAP rule computes $\sum_{i=0}^{N-1} r_i \cos(2\pi f_0 i)$ and compares it with $\frac{NA}{4}$ to decide $H_0$ or $H_1$. Now because of the phase uncertainty the received signal we have $r_i = A \cos(2\pi f_0 i + \theta) + Z_i$. 

2
Assuming that \( Z = \sum_{i=0}^{N-1} r_i \cos(2\pi f_0 i) \) as before, which is a Gaussian random variable, we compute its mean and variance under \( H_1 \).

\[
E[Z|H_1] = A \sum_{i=0}^{N-1} \cos(2\pi f_0 i) \cos(2\pi f_0 i + \theta_0)
\]

\[
= A \left( \sum_{i=0}^{N-1} \cos^2(2\pi f_0 i) \cos(\theta_0) - \sum_{i=0}^{N-1} \cos(2\pi f_0 i) \sin(2\pi f_0 i) \sin(\theta_0) \right)
\]

\[
\approx NA^2 \cos(\theta_0),
\]

where we used the trigonometric identity \( \cos(\alpha + \beta) = \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta) \).

\[
Var[Z] = \sigma^2 \sum_{i=0}^{N-1} \cos^2(2\pi f_0 i) = \frac{N\sigma^2}{2},
\]

which implies that :

\[
Z|_{H_1} \sim N \left( \frac{NA^2}{2} \cos(\theta_0), \frac{N\sigma^2}{2} \right).
\]

Hence we obtain the conditional error probability as

\[
P(E|H_1) = P \left( Z < \frac{NA}{4} \right)
\]

\[
= Q \left( \frac{NA^2 \cos(\theta_0) - \frac{NA}{4}}{\sqrt{\frac{N\sigma^2}{2}}} \right)
\]

\[
= Q \left( (2\cos(\theta_0) - 1) \sqrt{\frac{N\text{SNR}}{8}} \right)
\]

We see that if \( |\theta_0| > \frac{\pi}{3} \), \( P(E|H_1) > \frac{1}{2} \).

4. 5. The conditional error probability under \( H_0 \) does not depend on \( \theta_0 \) because under \( H_0 \) the transmitter does not send any signal and we receive pure noise. Hence

\[
P(E) = \frac{1}{2} \left\{ Q \left( \sqrt{\frac{N\text{SNR}}{8}} \right) + Q \left( (2\cos(\theta_0) - 1) \sqrt{\frac{N\text{SNR}}{8}} \right) \right\}
\]

We can simply see that in the case \( \theta_0 = \frac{\pi}{2} \):

\[
P(E) = \frac{1}{2} \left( Q \left( \sqrt{\frac{N\text{SNR}}{8}} \right) + Q \left( \sqrt{\frac{N\text{SNR}}{8}} \right) \right) = \frac{1}{2},
\]

3
where we used the identity \( Q(x) + Q(-x) = 1, \quad x \in \mathbb{R} \), and even worse, when \( \theta_0 = \pi \) (complete phase change),

\[
P(E) = \frac{1}{2} \left( Q \left( \sqrt{\frac{SNR}{8}} \right) + Q \left( -3\sqrt{\frac{SNR}{8}} \right) \right) > \frac{1}{2}.
\]

This shows that the phase uncertainty can completely disrupt the communication.

**Problem 2.** *(Gram-Schmidt Procedure On Tuples)*

We denote inner product by \(<,>\), norm by \(||.||\) intermediate vectors by \(\phi\) and final normalized vector by \(\psi\). We start from \(\beta_1\).

1. \(||\beta_1|| = \sqrt{<\beta_1,\beta_1>} = \sqrt{3}\). \(\psi_1 = \frac{\beta_1}{||\beta_1||} = \left(\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)\).

2. \(<\psi_1,\beta_2> = \sqrt{3}\). \(\phi_2 = \beta_2 - \sqrt{3}\psi_1 = (1, 1, -1, 0)\). \(||\phi_2|| = \sqrt{3}\) and so \(\psi_2 = \frac{\phi_2}{||\phi_2||} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, 0\right)\).

3. \(<\psi_1,\beta_3> = 0\) and \(<\psi_2,\beta_3> = 0\). \(\phi_3 = \beta_3 - 0\psi_1 - 0\psi_2 = (1, 0, 1, -2)\) and \(||\phi_3|| = \sqrt{1 + 1 + 4} = \sqrt{6}\) and so \(\psi_3 = \frac{\phi_3}{||\phi_3||} = (\frac{1}{\sqrt{6}}, 0, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}})\).

4. \(<\psi_1,\beta_4> = \sqrt{3}, <\psi_2,\beta_4> = 0\) and \(<\psi_3,\beta_4> = \sqrt{6}\). \(\phi_4 = \beta_4 - \sqrt{3}\psi_1 - 0\psi_2 - \sqrt{6}\psi_3 = (0, 0, 0, 0)\).

As can be seen the last vector is zero and this shows that the dimensionality of the space spanned by \(\beta_1, \cdots, \beta_4\) is only 3 not 4. So the other benefit of Gram-Schmidt orthogonalization is that it gives us the dimension of the space spanned by initial vectors.

**Problem 3.** *(Gram-Schmidt Procedure On Waveforms)*

An orthonormal basis may be found using the so-called Gram-Schmidt procedure.

1. We use Gram-Schmidt procedure:

   (a) The first step is to normalize the function \(s_1(t)\), i.e. the first function of the basis that we are looking for is

\[
\phi_0(t) = \frac{s_0(t)}{||s_0(t)||} = \frac{s_0(t)}{\sqrt{\int_0^1 s_0(t)^2 dt}} = \frac{s_0(t)}{\sqrt{\int_0^1 4t^2 dt}} = \frac{\sqrt{3}}{2}s_0(t) = \begin{cases} 0 & \text{if } t < 0 \\ \sqrt{3}t & \text{if } 0 < t < 1 \\ 0 & \text{if } t > 1 \end{cases}.
\]

   (b) Next, we subtract from \(s_1(t)\) the components that are in the span of the currently established part of the basis, i.e. in the span of \(\{\phi_0(t)\}\). This can be achieved by
projecting \( s_1(t) \) onto \( \phi_0(t) \) and then subtracting this projection from \( s_1(t) \), i.e.

\[
e_1(t) = s_1(t) - \langle s_1(t), \phi_0(t) \rangle \phi_0(t) = s_1(t) - \left( \int s_1(t) \phi_0(t) dt \right) \phi_0(t)
\]

\[
= s_1(t) - \left( \frac{\sqrt{3}}{2} \right) \left( \frac{4}{3} \right) \phi_0(t)
\]

\[
= s_1(t) - \frac{2}{\sqrt{3}} \phi_0(t)
\]

\[
= s_1(t) - s_0(t)
\]

From this, we find the second basis element as

\[
\phi_1(t) = \frac{e_1(t)}{||e_1(t)||} = \begin{cases} 
0 & \text{if } t < 1 \\
\sqrt{3} - \sqrt{3}(t - 1) & \text{if } 1 < t < 2 \\
0 & \text{if } t > 2
\end{cases}
\]

(c) Again, we subtract from \( s_2(t) \) the components that are in the span of the currently established part of the basis, i.e. in the span of \( \{\phi_0(t), \phi_1(t)\} \). This can be achieved by projecting \( s_2(t) \) onto \( \phi_0(t) \) and \( \phi_1(t) \) and then subtracting both these projections from \( s_2(t) \). For this step, it is essential that the basis elements \( \{\phi_0(t), \phi_1(t)\} \) be orthonormal. Make sure you understand why. Continuing the derivation, we obtain

\[
e_2(t) = s_2(t) - \langle s_2(t), \phi_0(t) \rangle \phi_0(t) - \langle s_1(t), \phi_1(t) \rangle \phi_1(t)
\]

\[
= s_2(t) - \left( \int s_2(t) \phi_0(t) dt \right) \phi_0(t) - \left( \int s_2(t) \phi_1(t) dt \right) \phi_1(t)
\]

\[
= s_2(t) - 0 - e_1(t)
\]

\[
= s_2(t) - s_1(t) + s_0(t),
\]

and from this, we find the third basis element as

\[
\phi_2(t) = \frac{e_2(t)}{||e_2(t)||} = \begin{cases} 
0 & \text{if } t < 2 \\
-\sqrt{3}(t - 2) & \text{if } 2 < t < 3 \\
0 & \text{if } t > 3
\end{cases}
\]

2. By definition we can write \( V_1(t) \) and \( V_2(t) \) as follows

\[
V_1(t) = 3\phi_0(t) - \phi_1(t) + \phi_2(t) = \begin{cases} 
3\sqrt{3}t & \text{if } 0 < t < 1 \\
-(\sqrt{3} - \sqrt{3}(t - 1)) & \text{if } 1 < t < 2 \\
-\sqrt{3}(t - 2) & \text{if } 2 < t < 3
\end{cases}
\]

and

\[
V_2(t) = -\phi_0(t) + 2\phi_1(t) + 3\phi_2(t) = \begin{cases} 
-\sqrt{3}t & \text{if } 0 < t < 1 \\
2(\sqrt{3} - \sqrt{3}(t - 1)) & \text{if } 1 < t < 2 \\
-3\sqrt{3}(t - 2) & \text{if } 2 < t < 3
\end{cases}
\]
3. We know that $V_1(t)$ and $V_2(t)$ are both real, thus
\[
\langle V_1(t), V_2(t) \rangle = \int V_1(t) V_2(t) dt = \langle V_1, V_2 \rangle = -3 * 1 - 1 * 2 + 1 * 3 = -2
\]

**Problem 4. (Matched Filter Intuition)**

1. The Cauchy-Schwarz inequality states
\[
|\langle x, y \rangle| \leq ||x|| ||y||
\]
with equality if and only if $x = \alpha y$ for some scalar $\alpha$. For our problem, we can write
\[
|\langle s, \phi \rangle|^2 \leq |s|^2 |\phi|^2 = |s|^2
\]
with equality if and only if $\phi = \alpha s$ for some scalar $\alpha$. Thus, the maximizing $\phi(t)$ is simply a scaled version of $s(t)$.

*Note:* In two dimensions, we have $|\langle x, y \rangle| \leq ||x|| ||y|| \cos \alpha$, where $\alpha$ is the angle between the two vectors; then, it is clear that the maximum is achieved when $\cos \alpha = 1$ $\iff \alpha = 0$ (or $\alpha = k2\pi$). Thus, $x$ and $y$ are colinear.

2. The inner product $\langle x, y \rangle$ is (using the definitions in the figure below) just the product of the length $x'$ and the length $y$, i.e. $\langle x, y \rangle = ||x'|| ||y||$. But it is immediately clear that $||x'||$ is maximal when $x$ points in the same direction as $y$.

3. Denote $s = (s_1, s_2)$ and $\phi = (\phi_1, \phi_2)$. The problem is
\[
\max_{\phi_1, \phi_2} (s_1 \phi_1 + s_2 \phi_2) \text{ subject to } \phi_1^2 + \phi_2^2 = 1.
\]
Thus, we can reduce by setting $\phi_2 = \sqrt{1 - \phi_1^2}$ to obtain

$$\max_{\phi_1} \left( s_1 \phi_1 + s_2 \sqrt{1 - \phi_1^2} \right).$$

This maximum is found by taking the derivative:

$$\frac{d}{d\phi_1} \left( s_1 \phi_1 + s_2 \sqrt{1 - \phi_1^2} \right) = s_1 - s_2 \frac{\phi_1}{\sqrt{1 - \phi_1^2}}.$$

Setting this equal to zero yields $s_1 = s_2 \frac{\phi_1}{\sqrt{1 - \phi_1^2}}$, i.e.

$$s_1^2 = s_2^2 \frac{\phi_1^2}{1 - \phi_1^2}.$$

This immediately gives $\phi_1 = \frac{s_1}{\sqrt{s_1^2 + s_2^2}}$, and thus $\phi_2 = \frac{s_2}{\sqrt{s_1^2 + s_2^2}}$, as expected.

Note: the goal of this exercise was to display yet another way to derive the matched filter.

4. Passing an input $s(t)$ through a filter with impulse response $h(t)$ generates output waveform $y(t) = \int s(\tau)h(t - \tau)d\tau$. If this waveform $y(t)$ is sampled at time $t = T$, then the output sample is:

$$y(T) = \int s(\tau)h(T - \tau)d\tau$$

(1)

An example signal $s(\tau)$ is shown in Figure 1(a). The filter is then the waveform shown in 1(b), and the convolution term of the filter in 1(c). Finally, the filter term $h(T - \tau)$ of Equation 1 is shown in 1(d). One can see that $h(T - \tau) = s(\tau)$, so indeed

$$y(T) = \int s(\tau)h(T - \tau)d\tau = \int s^2(\tau)d\tau = \int_0^T s^2(\tau)d\tau.$$

(v) Denote the signal spectrum by

$$S(f) = \int s(t)e^{-j2\pi ft}dt = |S(f)|e^{j\theta(f)}.$$

Then, the spectrum of the matched filter can be written as

$$H(f) = \int h(t)e^{-j2\pi ft}dt = \int s(T - t) e^{-j2\pi ft}dt = e^{-j2\pi fT}S^*(f) = e^{-j(\theta(f) + 2\pi fT)}|S(f)|,$$

where $S^*(f)$ is the complex conjugate. Now consider the signal $f(t)$ at the output of the matched filter. It is the convolution of the signal $s(t)$ with the matched filter
impulse response $h(t)$. As an inverse Fourier transform,

$$f(t) = \int S(f)H(f)e^{j2\pi ft} df$$

$$= \int |S(f)|^2e^{-j2\pi ft}e^{j2\pi ft} df$$

$$= \int |S(f)|^2e^{j2\pi ft-T} df.$$

Obviously, if $t = T$, all components in the integral “add in phase”, and we simply get

$$f(t = T) = \int |S(f)|^2 df.$$

**Problem 5. (AWGN Channel And Sufficient Statistic)**

For the first part we have:

1. Under hypothesis $H = i$, the received waveform is $Y(t) = s_i(t) + Z(t)$ and there is one-to-one correspondence between $Y(t)$ and $Y = (Y_0, Y_1, Y_2)^T$ where $Y_i = \langle Y(t), \phi_i(t) \rangle$. Hence, $Y$ is a sufficient statistic. It is straightforward to verify that when $H = i$, $Y = s_i + N$.

2. The third component of $s_i$ is zero for all $i$. Further more $N_0, N_1$ and $N_2$ are zero mean iid Gaussian random variables. Hence, $f_{Y|H}(y|i) = f_{N_0}(y_0 - s_{i0})f_{N_1}(y_1 - s_{i1})f_{N_2}(y_2)$ which is in the form $g_i(T(y))h(y)$ for $T(y) = (y_0, y_1)^T$ and $h(y) = f_{N_2}(y_2)$. Hence, by the Fisher-Neyman factorization theorem, $T(Y) = (Y_0, Y_1)^T$ is a sufficient statistic.
For the second part of the problem in which $N_2 = N_1$ we have:

1. If we have only $(Y_0, Y_1)$ then hypothesis testing problem will be $H = i : (Y_0, Y_1) = (s_{i0}, s_{i1}) + (N_0, N_1)$ $i = 0, 1$. Using the fact that $s_0 = (1, 0, 0)^T$ and $s_1 = (0, 1, 0)^T$, the ML test becomes $y_0 - y_1 > 0$. Under $H = 0$, $Y_0 - Y_1$ is a Gaussian random variable with mean 1 and variance $2\sigma^2$ and so $P_e(0) = Q\left(\frac{1}{\sqrt{2\sigma^2}}\right)$. By symmetry $P_e(1) = Q\left(\frac{1}{\sqrt{2\sigma^2}}\right)$ and so the probability of the error will be $P_e = \frac{1}{2}(P_e(0) + P_e(1)) = Q\left(\frac{1}{\sqrt{2\sigma^2}}\right)$.

2. Now assume that we have access to $Y_0$, $Y_1$ and $Y_2$. $Y_2$ contains $N_2$ under both hypotheses. Hence, $Y_1 - Y_2 = s_{i1} + N_1 - N_2 = s_{i1}$. This shows that at the receiver we can observe the second component of $s_i$ without noise. As the second component is different under both hypotheses, we can make an error-free decision about $H$ and the decision rule will be:

   $$\hat{H} = \begin{cases} 
   0 & y_1 - y_2 = 0 \\
   1 & y_1 - y_2 = 1 
   \end{cases}$$

   Clearly this decision rule minimizes the probability of the error.

3. $Y_2$ allows us to reduce the probability of the error. Hence, $(Y_0, Y_1)$ can’t be a sufficient statistic.